J Chron Dis Vol. 40, Suppl. 2, pp. 171S–179S, 1987
Printed in Great Britain

0021-9681/87 $3.00 + 0.00
Pergamon Journals Ltd

USE OF MULTISTAGE MODELS TO INFER STAGE AFFECTED BY CARCINOGENIC EXPOSURE: EXAMPLE OF LUNG CANCER AND CIGARETTE SMOKING

CHARLES C. BROWN and KENNETH C. CHU

Division of Cancer Prevention and Control, National Cancer Institute, Bethesda, MD 20205, U.S.A.

Abstract—Statistical methods based on the Armitage–Doll mathematical model of the carcinogenic process are presented for analyzing epidemiologic case–control studies of cancer. These methods are proposed to provide inferences regarding the stage(s) in the cancer process at which the exposure of interest acts. An example of these methods is given which shows evidence that carcinogens in cigarette smoke appear to affect the transition rates for two separate stages in the development of lung cancer, and the relative magnitudes of these effects are estimated. The data for this analysis came from a European multi-center case–control study of lung cancer.

The results of the analysis show that: (1) the relative risk of lung cancer among continuing smokers compared to nonsmokers of the same age decreases as the age started smoking increases, while the rate of smoking stays fixed, a result which indicates a carcinogenic effect on an early stage in the process; and (2) the relative risk among ex-smokers compared to continuing smokers having the same duration and rate of smoking decreases with time since smoking stopped, a result which indicates a carcinogenic effect on a late stage in the process. Both results are shown to be best described by the hypothesis that cigarette smoking affects two stages. The estimated relative magnitudes of cigarettes' carcinogenic effects on the two stages indicate that the largest proportion of the total lifetime lung cancer risk among continuing smokers is due to its late stage effect, and that the proportion of risk due to causes other than smoking varies from 23% among men smoking 1–10 cigarettes per day to 6% among those smoking greater than 30 cigarettes per day. These findings imply that preventive measures directed toward inducing smokers to stop would have a potentially substantial payoff in reducing future lung cancer mortality.

1. INTRODUCTION

The results of epidemiologic studies of human cancer are becoming more frequently interpreted in terms of quantitative multistage theories of the carcinogenic process. Mathematical models derived from these theories were originally proposed to explain the observation that age-specific cancer incidence and mortality rates for many human cancers of epithelial origin increase with the fifth or sixth power of age (Muller [1], Nordling [2]). The age pattern of occurrence of many tumors induced in experimental animals by continuous exposure to chemical carcinogens is also predicted by these multistage theories. Experimental regimens involving initiation and promotion provide direct evidence of the effect of early and late stage events in the carcinogenic process (Stenback *et al.* [3–5]). In more recent years, the models are being used to elucidate the meaning of the relationships observed between cancer risk and age at exposure, duration of exposure, and time since exposure ended (Whittemore [6], Day and Brown [7], Brown and Chu [8]). In this paper we shall show how particular time patterns of cancer risk may be used to infer the stage in the cancer process, early or late, at which the carcinogen exerts its influence. To demonstrate the type of analyses we propose, we present an example of the relationship between cigarette smoking and lung cancer for which we shall estimate the relative contributions of smoking upon early and late stages in the development of lung cancer. We shall limit our presentation to inferences derived from the mathematical model of carcinogenesis proposed by Armitage and Doll [9].

PM3003530417

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

2. EXAMPLE OF CIGARETTE SMOKING AND LUNG CANCER

For a number of years cigarette smoking has been known to be associated with an increased risk of developing lung cancer [10]. Epidemiologic studies have shown that the risk of lung cancer among continuing smokers increases at a high power of the duration of smoking. These studies have also shown that once an individual stops smoking, his excess lung cancer risk ceases to increase and remains nearly frozen at the level attained when he last smoked. Doll [11], among others, has interpreted these findings as implying that smoking affects two stages, one early and one late, in a presumed multistage carcinogenic process. The observation that the absolute excess age-specific lung cancer risk among smokers increases with the same power of duration of smoking as does the risk with age among nonsmokers indicates that smoking likely affects the first stage of the process. On the other hand, the rapidity of the effect of stopping smoking points to smoking's effect at a late stage. In addition, Doll and Peto [12] have determined that the dose–response between dose rate measured in cigarettes per day and response measured as excess lung cancer incidence rate appears to be quadratic, which they interpreted to imply that smoking affects two stages.

Though there is general agreement that smoking likely affects both an early and a late stage in the process of lung cancer development, no one has quantified the relative magnitude of smoking's effects upon the two stages. The magnitude of these effects could have substantial importance in the planning and study of intervention methods to reduce lung cancer incidence. If smoking is primarily an early stage carcinogen, then the early smoking years will be the primary cause of future lung cancer and intervention activities directed at teenagers who have not yet started to smoke or who have been smoking for a short while would be expected to have the greatest payoff. However, if smoking primarily affects a late stage, then an intervention designed to get people to stop smoking at all ages will definitely have an effect on their future risk. The remainder of this section describes our methods for estimating these relative effects of smoking on early and late stages in lung cancer development.

The data we use are from a hospital-based case–control interview study of individuals newly diagnosed with lung cancer in five Western European countries. Patients admitted to a study hospital with histologically confirmed lung cancer during 1976–80 were entered into the study. Two age-, sex- and centre-matched controls were selected for each patient. The controls were hospital patients in whom a non-tobacco related disease had been diagnosed. No surrogate interviews were employed. Additional details of the study design are described by Lubin *et al.* [13]. That part of the study which we shall analyze here consists of 6920 male patients and 13,460 male controls. Female patients and controls comprise only 11% of the total and thus we have simplified the analysis by limiting ourselves to the results for males. We classify as smokers only those individuals who have ever smoked cigarettes.

Our method of analysis is to examine the time and age patterns of the relative risk of lung cancer to quantify the effects of cigarette smoking upon presumed early and late stages in the cancer process. Following Day and Brown [7] and Doll and Peto [12] we assume that cigarette smoking increases both the rates of first stage and penultimate stage cellular events. Our analysis is divided into two separate comparisons: (1) examination of the manner in which relative risk varies with age started smoking among continuing smokers; and (2) examination of the manner in which relative risk varies with time since stopped smoking among ex-smokers. We have performed these two separate analyses because of the apparently conflicting epidemiologic evidence among continuing smokers compared to nonsmokers which points to smoking affecting an early stage and the evidence from ex-smokers compared to continuing smokers which points to smoking's effect on a late stage. We shall show that the two separate analyses are consistent with the hypothesis that cigarette smoking affects two stages.

To estimate these effects of cigarette smoking on the first and penultimate stages in the development of lung cancer, we use the relative risk models given in the Appendix which are derived from the results in Day and Brown [7]. Our analysis consists of two steps. The first step is to use unconditional logistic regression to estimate the relative risks for the different categories of the time/age variables of interest. The second step was to fit the relative risk models in Appendix equations (A1) and (A2) to these relative risk estimates to derive estimates for the parameters r_1 and r_{k-1}. The relative magnitude of

PM3003530418

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

these parameters indicates the degree to which each stage of the cancer process is affected by the carcinogenic exposure.

We use two separate logistic regression models to estimate the relative risks of interest. The first logistic analysis estimates the effect of age started smoking upon the relative risk of lung cancer among continuing smokers compared to nonsmokers. Ex-smokers were not used in this analysis. The model for this comparison contained categorical variables for the matching variables study area (seven categories) and age at interview (10 categories from age $\leqslant 46$ through age $\geqslant 73$), and for the factors number of cigarettes smoked per day (four categories, 1–10 as the referent category, 11–20, 21–30 and $\geqslant 31$), frequency of inhalation (all the time vs less as the referent category), percent of total duration smoking nonfiltered cigarettes (all the time vs less as the referent category), and the variable of primary interest, age started smoking (eight categories from $\leqslant 14$, 15, 16, 17, 18, 19–20, $\geqslant 21$, with nonsmokers as the referent category). These particular age at interview and age started smoking categories were selected to produce, as nearly as possible, equal numbers of individuals in each category. The estimated logistic regression coefficients for the variables age started smoking, number of cigarettes smoked per day, frequency of inhalation and percent of time smoking nonfiltered cigarettes are given in Table 1. As expected, the relative risk increases with the number smoked per day; men who smoke at least $1\frac{1}{2}$ packs per day have 3 times the risk of men who smoke $\frac{1}{2}$ pack or less each day. Except for those who started smoking at age 14 or younger, the relative risk steadily declines with increasing age started smoking. We examined without success the data for other variables which might explain our finding a lower relative risk than expected for those who started smoking at age 14 or younger; this age group had a similar distribution of other risk factors as did the other groups. The decrease in relative risk with increasing age started exposure adjusted for age at diagnosis and amount smoked per day implies that smoking affects an early stage in the carcinogenic process since Appendix equation (A1) with $r_1 = 0$ shows that the relative risk would be nearly independent of age started smoking if smoking affected only a late stage. However, the moderate magnitude of this decrease is not consistent with the hypothesis that smoking affects only the first stage, for if this was the case the decrease would be much more pronounced. This is shown graphically in Fig. 1 which compares the observed pattern of relative risk with patterns predicted by three particular hypotheses. The predicted patterns are obtained from least squares fits of the logarithm of equation (A1) to the estimated log relative risks shown in Table 1. The three models assumed (a) only a late stage affect, $r_1 = 0$, (b) only an early stage affect, $r_{k-1} = 0$, and (c) an affect on both stages. Different values of k and t were tried, with $k = 5$ and $t = 65$ giving the best overall fits. This figure shows that the hypothesis of two affected stages clearly fits the

Table 1. Estimated logistic regression coefficients in continuing cigarette smokers compared to nonsmokers

Variable	Estimated coefficient	Relative risk
Age started smoking		
Nonsmoker		1.0
$\leqslant 14$	1.288 ($\pm$0.101)	3.6
15	1.419 ($\pm$0.106)	4.1
16	1.390 ($\pm$0.106)	4.0
17	1.384 ($\pm$0.110)	4.0
18	1.294 ($\pm$0.097)	3.6
19–20	1.225 ($\pm$0.094)	3.4
$\geqslant 21$	1.077 ($\pm$0.094)	2.9
Number cigarettes smoked per day		
1–10		1.0
11–20	0.712 ($\pm$0.059)	2.0
21–30	1.035 ($\pm$0.070)	2.8
$\geqslant 31$	1.184 ($\pm$0.074)	3.3
Frequency of inhalation		
All the time	0.117 ($\pm$0.046)	1.1
Percent of time smoked nonfiltered cigarettes		
All the time	0.179 ($\pm$0.049)	1.2

PM3003530419

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

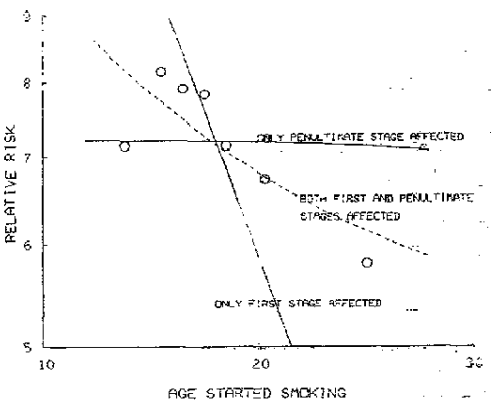

Fig. 1. Relative risks of lung cancer in smokers vs nonsmokers of the same age estimated from logistic regression model: observed pattern compared to patterns predicted by multistage model of carcinogenesis.

observed pattern substantially better than either of the separate single stage hypotheses.

A separate logistic model estimates the relationship between relative risk and the time since smoking stopped using continuing smokers as the referent group. Brown and Chu [8] showed that duration of exposure must be controlled for to make valid use of the relative risk pattern as an indication of which carcinogenic stage is affected by exposure. Therefore, we restrict our analysis only to individuals who have ever been cigarette smokers. It must be kept in mind that individuals who give up the smoking habit do so for a variety of reasons, some related to health. In a preliminary analysis we found that the lung cancer risk relative to continuing smokers was greater than unity for men who had recently quit smoking. The relative risk was a maximum for those who had quit within one year of the interview. Those individuals who had quit smoking were asked in the questionnaire to identify the reason for quitting as being health related or not. Figure 2 shows the relationship between the years since quitting and the proportion of men who answered that they quit for health reasons. This figure clearly shows that men who had recently quit did so more often for health reasons than men who had quit a longer time in the past. Since we felt this association might bias the relationship between relative risk and time since stopped smoking, we decided to include those who quit within the past year with the continuing smokers, eliminate from the analysis the 420 men who quit between one and two years before the interview, and identify as ex-smokers only those men who had quit for more than two years. In addition, we included the reason for quitting as an adjustment variable in our logistic regression. The regression model also contained categorial variables for study area, age at interview, number of cigarettes smoked, duration of smoking (nine categories in five year groups beginning with $\leqslant 10$ years as the referent group), frequency of inhalation, percent of time smoking nonfiltered cigarettes, and years since stopped smoking (10 categories from continuing smokers as the referent group to 3, 4, 5–6, 7–8, 9–11, 12–15, 16–20, 21–26 and $\geqslant 27$ years since stopped smoking). The factor age started smoking, shown to be important in the previous analysis, was not included here because of its multicolinearity with the other factors, current age, duration of smoking, and time since quit smoking. As for the other logistic analysis, these particular years since smoking stopped categories were selected to produce near equal numbers of individuals in each. Table 2 gives the logistic regression coefficient estimates for duration of smoking, number of cigarettes smoked per day, frequency of inhalation, percent of time smoking nonfiltered cigarettes, reason for stopping smoking, and time since stopped smoking. The relative risk of lung cancer increases with duration of smoking as predicted by the multistage theory, and decreases with increasing time since smoking stopped. In addition, the risk ratio for an individual who stopped for health reasons relative to an individual who stated that he stopped for reasons other than health was estimated to be 1.3 ($p < 0.001$). As in the previous analysis the relative risk increased with the number of cigarettes smoked per day. The relative risk estimates are 1.0, 1.9, 2.4 and 2.8 which are nearly equal to those shown in Table 1.

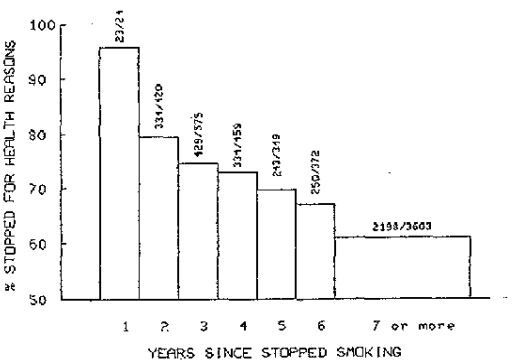

Fig. 2. Proportion of ex-smokers who stopped for reasons related to health.

The risk
relative to
the same
decreases
to level o
after abo
ically in
compared
hypothese
smoking.
three lea
Appendix
estimates
the first
rejected
would b
exposure
clearly d
function
These d
that only
carcinog
affected

PM3003530420

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

Table 2. Estimated logistic regression coefficients in ex-smokers compared to continuing smokers

Variable	Estimated coefficient	Relative risk
Years since stopped smoking		
0–1		1.0
3	−0.014 (±0.098)	0.99
4	−0.248 (±0.106)	0.78
5–6	−0.346 (±0.090)	0.71
7–8	−0.375 (±0.102)	0.69
9–11	−0.724 (±0.095)	0.48
12–15	−0.751 (±0.107)	0.47
16–20	−0.939 (±0.116)	0.39
21–26	−0.828 (±0.120)	0.44
⩾27	−0.907 (±0.137)	0.40
Years of smoking		
⩽10		1.0
11–15	−0.001 (±0.224)	1.0
16–20	0.527 (±0.190)	1.7
21–25	0.750 (±0.180)	2.1
26–30	0.954 (±0.175)	2.6
31–35	1.083 (±0.173)	3.0
36–40	1.235 (±0.173)	3.4
41–45	1.338 (±0.174)	3.8
⩾46	1.512 (±0.173)	4.5
Number of cigarettes smoked per day		
1–10		1.0
11–20	0.642 (±0.049)	1.9
21–30	0.878 (±0.057)	2.4
⩾31	1.032 (±0.060)	2.8
Frequency of inhalation		
All the time	0.127 (±0.039)	1.1
Percent of time smoked nonfiltered cigarettes		
All the time	0.101 (±0.040)	1.1
Reason for stopping smoking		
Health-related	0.241 (±0.064)	1.3

The risk of lung cancer among ex-smokers relative to that among continuing smokers of the same age and smoking rate and duration decreases with time since quitting, and appears to level out and possibly even begin to increase after about 20 years. This is displayed graphically in Fig. 3 where the observed pattern is compared to patterns predicted by different hypotheses concerning the stage(s) affected by smoking. These predicted patterns are from three least squares fits of the logarithm of Appendix equation (A2) to the log relative risk estimates in Table 2. The hypothesis that only the first stage is affected by smoking is clearly rejected by these data since the relative risk would be an increasing function of time since exposure stopped and the observed pattern is clearly decreasing (the "best" fitting increasing function is nearly flat as shown in the figure). These data are consistent with the hypothesis that only the penultimate stage is affected by the carcinogenic exposure. However, the single affected penultimate stage hypothesis does not predict the upturn in relative risk which is suggested by these data. A hypothesis that assumes smoking affects both the first and penultimate stages provides a better description,

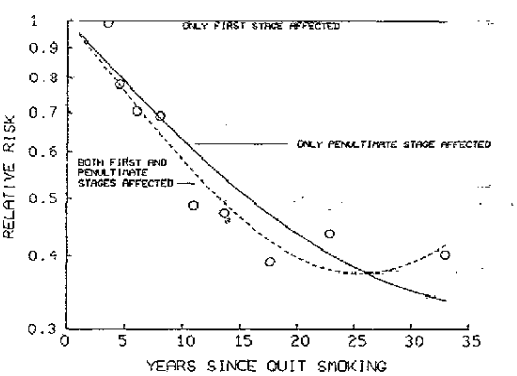

Fig. 3. Relative risks of lung cancer in ex-smokers vs continuing smokers of the same age and same smoking duration estimated from logistic regression model: observed pattern compared to patterns predicted by multistage model of carcinogenesis.

PM3003530421

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

Table 3. Estimated relative increase in cellular event rates for multistage model applied to relative risks of lung cancer among cigarette smokers

Number cigarettes smoked per day	Relative increase caused by smoking	
	First stage (r_1)	Penultimate stage (r_{k-1})
1–10	0.7	2.8
11–20	2.5	5.0
21–30	3.5	6.3
⩾31	4.0	7.0

in terms of a smaller sum of squares, of the observed pattern since it does predict an upturn in relative risk. This delay period before the upturn can be thought of as representing the time required for an initiated cell, for which the first cellular change was caused by its exposure to carcinogens in cigarette smoke, to pass spontaneously through the remaining stages, including the penultimate in the carcinogenic process.

Both analyses, the relationship of lung cancer relative risk to age started exposure to cigarettes, and the relationship of risk to time since exposure stopped, provide consistent evidence for the two affected stage hypothesis. To quantify the magnitude of smoking's effect on the two stages, the logarithm of the relative risk functions in Appendix equations (A1) and (A2) were simultaneously fit to the log relative risk estimates derived from the two sets of logistic regression parameter estimates given in Tables 1 and 2. This was done for each category of number of cigarettes smoked daily to estimate r_1 and r_{k-1} as a function of number of cigarettes smoked per day. The relative risks for age started smoking given in Table 1 relate to the referent category of 1–10 cigarettes smoked per day. Since the logistic regression model implicitly assumes a multiplicative relation between age started smoking and number of cigarettes smoked per day, the set of relative risk estimates for individuals who smoke more than 1–10 per day is given by the product of the appropriate relative risks (e.g. the relative risk estimate for an individual who started smoking at age 16 and smoked 21–30 cigarettes per day is $4.0 \times 2.8 = 11.2$). Since the relative risk estimates derived from the age started smoking coefficients in Table 1 relate to the referent categories of the inhalation and nonfiltered cigarette factors, we adjusted the log relative risks to reflect the actual distribution of continuing smokers who inhaled all the time (69%) and who smoked only nonfiltered cigarettes (27%). This resulted in adding 0.13 to each log relative risk.

A nonlinear least squares procedure was used to estimate the unknown relative risk parameters, r_1 and r_{k-1}. The estimated values of r_1 and r_{k-1} for each of the four fits are given in Table 3. These values are based on $k = 5$, $t = 65$ and $d = 25$ for each of the four least square fits. We found these values to produce the smallest overall sum of squares when totalled over the four fits. These estimates in Table 3 imply that cigarette smoking produces a larger relative increase on a late stage event in the carcinogenic process than on the event which initiates it. For an individual who smokes 21–30 cigarettes per day, this carcinogenic exposure is estimated to increase the rate of the early cellular event by 350% over background and the rate of the late cellular event by 630%.

Taken by themselves, these estimated relative increases in cellular event rates are difficult to interpret. However, in the Appendix we present an approach to measure their contributions to

Table 4. Relative lifetime* risk of lung cancer among continuing smokers who begin at age 17

Number cigarettes smoked per day	Lifetime risk contribution due to smoking-induced effect			
	Background	Only first stage	Only penultimate stage	Both first and penultimate stages
1–10	0.23	0.04	0.63	0.10
11–20	0.11	0.06	0.53	0.30
21–30	0.08	0.06	0.47	0.39
⩾31	0.06	0.06	0.45	0.43

*Through age 85.

the lifetime ri
divides this ri
ground risk d
the excess risk
the first cellul
cellular event,
imate events.
4. These calcu
the median ag
in this study.
inverse relatic
rettes smoke
background
23% for "ligl
6% for "hea
magnitude o
with estimat
lung cancer
butable to ci
the largest p
order of 50%
effect on onl
5% of the t
on only the
increases, th
lar events b
The proport
effect on bot
increases fr
for "heavy"

The logi
analysis di
We searche

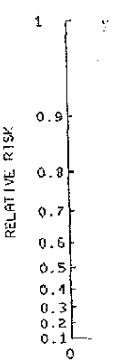

Fig. 4. Relat
to continuin
smoking est
adjusted fo

PM3003530422

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

the lifetime risk of lung cancer. This method divides this risk into four pieces: (1) the background risk due to causes other than smoking; the excess risk due to smoking causing (2) only the first cellular event, (3) only the penultimate cellular event, and (4) both the first and penultimate events. The results are presented in Table 4. These calculations assume $t_0 = 17$ which was the median age started smoking for the subjects in this study. As expected, this table shows an inverse relationship between the number of cigarettes smoked per day and the contribution of background to total lifetime risk ranging from 23% for "light" smokers (1–10 per day) to only 6% for "heavy" smokers (≥ 31 per day). The magnitude of these percentages agree closely with estimates of 85–90% for the amount of lung cancer estimated by others to be attributable to cigarette smoking [15]. In addition, the largest proportion of the total risk, on the order of 50%, is estimated to be from smoking's effect on only the penultimate stage while only 5% of the total risk is due to smoking's effect on only the first stage. As the exposure rate increases, the relative contribution of two cellular events being caused by smoking increases. The proportion of lifetime risk due to smoking's effect on both early and late stage cellular events increases from 10% for "light" smokers to 43% for "heavy" smokers.

3. DISCUSSION

The logistic regression models used in our analysis did not involve any interactive effects. We searched for the occurrence of such interactions which would produce different relative risk patterns than observed, but we did not find any. The relative risk patterns shown in Figs 1 and 2 are not changed to any substantial degree when we limit the analysis to include only individuals who smoked a particular number of cigarettes per day (e.g. 1–20 or ≥ 21) or to those cases and controls who were at particular ages at interview (e.g. ≤ 55 or > 65). We also examined the pattern of relative risk with time since stopped smoking for individuals with different smoking durations (e.g. 21–30 or 31–40 years) and found similar results. Thus the general patterns of relative risks seen in Figs 1 and 2 are consistently found across various strata of individuals and there appear to be no interactions with other factors. Other smoking characteristics were also examined to see if they should be included in the logistic regression models. However, we found that inhalation frequency and percent of time smoking nonfiltered cigarettes were the only statistically significant factors.

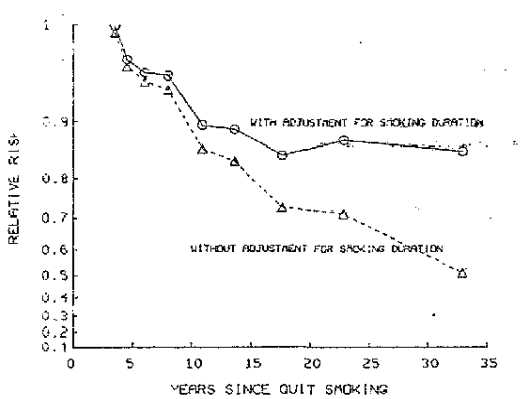

Fig. 4. Relative risks of lung cancer in ex-smokers compared to continuing smokers as a function of time since stopped smoking estimated from logistic regression model; pattern adjusted for smoking duration compared to pattern unadjusted for duration.

We wish to emphasize the importance of controlling duration of exposure when studying the pattern of relative risk as a function of time since exposure ended. Doll and Peto [12] found duration of smoking to be the single most important factor in the evolution of lung cancer risk. Therefore, an association of duration with other factors being studied could produce a substantial bias. In this study, the correlation between the duration of smoking and the time since smoking stopped for ex-smokers is -0.6 indicating that the men who had stopped smoking for many years had smoked for less time than those men who had stopped for a shorter time. Therefore, without adjusting for duration of exposure, a decreasing trend in relative risk with increasing time since exposure stopped would not necessarily indicate a late stage carcinogenic effect since the observed decrease might have been caused by the associated decrease in exposure duration. As an example of the potential for bias, we re-estimated the relative risks for the years-since-smoking-stopped categories without adjusting for smoking duration. Whereas the decreasing trend of the relative risks adjusted for duration shown in Table 2 flattened out as the years since smoking stopped increased, the decreasing trend continued sharply down when no adjustment as made for duration of smoking. This is shown in Fig. 4. For the category "≥ 27 years since stopped smoking", the relative risk estimate is 0.40 when duration adjustment is made (Table

PM3003530423

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

2) and is 0.17 when no adjustment is made. In this situation the bias produced by smoking duration is substantial and could lead to different conclusions.

The primary purpose of our analysis is to estimate the relative effects of cigarette smoking upon early and late events in the development of lung cancer. The results in Tables 3 and 4 indicate that smoking has a substantially greater effect on late stage events than it does on initiating events. This implies that preventive measures directed toward inducing smokers to stop should have a potentially substantial payoff in reducing the future burden of lung cancer mortality. However, this does not imply that one should ignore interventions designed to stop people from ever starting to smoke; since the duration of smoking is the single most important factor in the evolution of lung cancer, the future lung cancer mortality in the population will be best reduced by reducing the number of individuals who ever begin to smoke.

REFERENCES

1. Muller HG. Radiation damage to the genetic material. **Sci Prog** 1951; 7: 93–493.
2. Nordling CO. A new theory of the cancer inducing mechanism. **Br J Cancer** 1953; 7: 68–72.
3. Stenback F, Peto R, Shubik P. Initiation and promotion at different ages and doses in 2200 mice. I. Methods, and the apparent persistence of initiated cells. **Br J Cancer** 1981; 44: 1–14.
4. Stenback F, Peto R, Shubik P. Initiation and promotion at different ages and doses in 2200 mice. II. Decrease in promotion by TPA with aging. **Br J Cancer** 1981; 44: 15–21.
5. Stenback F, Peto R, Shubik P. Initiation and promotion at different ages and doses. III. Linear extrapolation from high doses may underestimate low-dose tumour risks. **Br J Cancer** 1981; 44: 24–33.
6. Whittemore AS. The age distribution of human cancer for carcinogenic exposures of varying intensity. **Am J Epidemiol** 1977; 106: 418–432.
7. Day NE, Brown CC. Multistage models and primary prevention of cancer. **J Natl Cancer Inst** 1980; 64: 977–989.
8. Brown CC, Chu KC. Implications of the multistage theory of carcinogenesis applied to occupational arsenic exposure. **J Natl Cancer Inst** 1983; 70: 455–463.
9. Armitage P, Doll R. The age distribution of cancer and a multistage theory of carcinogenesis. **Br J Cancer** 1954; 8: 1–12.
10. United States Public Health Service. **Smoking and Health.** Report of the Advisory Committee to the Surgeon General of the Public Health Service. United States Department of Health, Education and Welfare, Center for Disease Control: 1964.
11. Doll R. An epidemiologic perspective of the biology of cancer. **Cancer Res** 1978; 38: 3573–3583.
12. Doll R, Peto R. Cigarette smoking and bronchial carcinoma: dose and time relationships among regular smokers and life long non-smokers. **J Epidemiol Commun Health** 1978; 32: 303–313.
13. Lubin JH, Blot WJ, Berrino F, Flamant R, Gillis CR, Kunze M, Schmahl D, Visco G. Modifying risk of developing lung cancer by changing habits of cigarette smoking. **Br Med J** 1984; 288: 1953–1956.
14. U.S. Department of Health, Education and Welfare. **Vital Statistics of the United States 1970,** Vol. II: **Mortality Part A.** Rockville, Maryland: National Center for Health Statistics: 1975.
15. Loeb LA, Ernster VL, Warner KE, Abbotts K, Laszlo J. Smoking and lung cancer: an overview. **Cancer Res** 1984; 44: 5940–5958.

APPENDIX

Assuming that cigarette smoking affects both the first and penultimate stages in a multistage process of lung cancer development, relative risk models follow directly from Day and Brown [7, equation (4)]. We let t denote the individual's current age, t_0 the age at which smoking began, d the duration of smoking, and f the follow-up time since smoking stopped. As a function of age started smoking t_0, the relative risk of lung cancer in continuing smokers vs non-smokers of the same age t is,

$$\mathrm{RR}_c(t_0) = \frac{t^{k-1} + r_1(t-t_0)^{k-1} + r_{k-1}(t^{k-1} - t_0^{k-1}) + r_1 r_{k-1}(t-t_0)^{k-1}}{t^{k-1}}, \quad \text{(A1)}$$

where r_1 and r_{k-1} denote the relative increases in the first and penultimate cellular events produced by exposure to cigarette smoke. By relative increase we mean the following: let u denote the cellular event rate in the absence of smoking; let $u + v$ denote the rate when exposed to cigarette smoke; then $r = (v/u)$ represents the relative increase in the rate due to cigarette smoking. In a similar derivation, as a function of the time elapsed since smoking stopped f, the relative risk of lung cancer in ex-smokers vs continuing smokers of the same age t and the same duration of smoking d is given by,

$$\mathrm{RR}_x(f) = \frac{t^{k-1} + r_1[(d+f)^{k-1} - f^{k-1}] + r_{k-1}[(t-f)^{k-1} - (t-d-f)^{k-1}] + r_1 r_{k-1} d^{k-1}}{t^{k-1} + r_1 d^{k-1} + r_{k-1}[t^{k-1} - (t-d)^{k-1}] + r_1 r_{k-1} d^{k-1}}. \quad \text{(A2)}$$

For an individual exposed continuously beginning at age t_0, without stopping to a carcinogen which affects both the first and penultimate stages in the multistage process, the Armitage–Doll model of carcinogenesis leads to the following age-specific cancer incidence rates at age t,

$$\lambda_0(t) = \theta t^{k-1},$$

$$\lambda_1(t) = \theta r_1 (t-t_0)^{k-1}, \quad t \geq t_0,$$

$$\lambda_{k-1}(t) = \theta r_{k-1}(t^{k-1} - t_0^{k-1}), \quad t \geq t_0,$$

and

$$\lambda_{1,k-1}(t) = \theta r_1 r_{k-1}(t-t_0)^{k-1}, \quad t \geq t_0, \quad \text{(A3)}$$

where θ is a scale factor.

The function $\lambda_0(t)$ represents the spontaneous, or background rate; $\lambda_1(t)$ represents the excess rate due to a carcinogenic effect *only on the first stage* (e.g. the excess rate produced by smoking causing the first cellular event but not the penultimate event); $\lambda_{k-1}(t)$ represents the excess rate due to a carcinogenic effect on *only on the penultimate stage*; and $\lambda_{1,k-1}(t)$ represents the excess rate due to carcinogenic effects on *both the first and penultimate stages*. The total age-specific risk is the sum of the background and the excess,

$$\lambda(t) = \lambda_0(t) + \lambda_1(t) + \lambda_{k-1}(t) + \lambda_{1,k-1}(t). \quad \text{(A4)}$$

The lifetime risk can be estimated from competing risk

considerations as,

$$I$$

where $S(t)$ is the su
alive at age t. This li
pieces by replacing λ
parts given in equati
does not provide an
use the results of H
absolute magnitude
of lung cancer. Ho
contributions of the i
dividing each of the
factor cancels out).

$$L_0 =$$

$$L_1 =$$

PM3003530424

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001

considerations as,

$$L = \int_0^\infty \lambda(t)S(t)\,dt, \qquad \text{(A5)}$$

where $S(t)$ is the survival probability that an individual is alive at age t. This lifetime risk L can be factored into four pieces by replacing $\lambda(t)$ in equation (A5) by its component parts given in equation (A3). Since our case-control study does not provide an estimate of the scale factor θ, we cannot use the results of this case–control study to estimate the absolute magnitude of smoking's effect on the lifetime risk of lung cancer. However, we may estimate the relative contributions of the four components on the lifetime risk by dividing each of the four by their total (the unknown scale factor cancels out). These relative lifetime risks are,

$$L_0 = \int_0^\infty \lambda_0(t)S(t)\,dt/L,$$

$$L_1 = \int_0^\infty \lambda_1(t)S(t)\,dt/L,$$

$$L_{k-1} = \int_0^\infty \lambda_{k-1}(t)S(t)\,dt/L,$$

$$L_{1,k-1} = \int_0^\infty \lambda_{1,k-1}(t)S(t)\,dt/L. \qquad \text{(A6)}$$

Thus, for example, L_1 represents that proportion of the lifetime risk of lung cancer produced by smoking causing only the first stage of the carcinogenic process (the other stages being caused by "background"). We estimate these relative lifetime risks in equation (A6) by approximating the integrals by summations over single years of age up to age 85 where $S(t)$ represents U.S. national life table survival probabilities for white males [14] and the parameter estimates $\hat{r}_1$ and $\hat{r}_{k-1}$ are from Table 3. For this example we assume that the estimates $\hat{r}_1$ and $\hat{r}_{k-1}$, based on a European study, are applicable to U.S. males and we use the total U.S. male survival rates to approximate the survival rates for male smokers. When computing lifetime cancer risks, this approximation will produce overestimates since smokers do not live as long as nonsmokers. However, our computation of proportions of lifetime cancer risk should not be as biased by this use of total U.S. male survival rates.

PM3003530425

Source: https://www.industrydocuments.ucsf.edu/docs/hlnj0001